Discovery of Selective Bioactive Small Molecules by Targeting an RNA Dynamic Ensemble

Andrew C. Stelzer[1,4,5], **Aaron T. Frank**[2,5], **Jeremy D. Kratz**[1], **Michael D. Swanson**[3], **Marta J. Gonzalez-Hernandez**[3], **Janghyun Lee**[1], **Ioan Andricioaei**[2], **David M. Markovitz**[3], and **Hashim M. Al-Hashimi**[1,*]

[1]Department of Chemistry & Biophysics, University of Michigan, 930 North University Avenue, Ann Arbor, Michigan 48109, USA

[2]Department of Chemistry, University of California Irvine, 1102 Natural Sciences 2, Irvine, California 92697, USA

[3]Department of Internal Medicine, Division of Infectious Diseases, University of Michigan Medical Center, Ann Arbor, Michigan 48109, USA

Abstract

RNA is growing in its importance as a drug target but current approaches used to identify protein-targeting small molecules are ill-suited for RNA. By docking small molecules onto an RNA dynamic ensemble constructed by combining Nuclear Magnetic Resonance (NMR) spectroscopy and computational molecular dynamics, we virtually screened small molecules that target the entire structure landscape of the transactivation response element (TAR) from the human immunodeficiency type 1 virus (HIV-1). We quantitatively predict binding energies for small molecules that bind different RNA conformations and report the *de novo* discovery of six compounds that bind TAR with near record affinity and inhibit its interaction with a Tat peptide *in vitro* (K_is = 710 nM–169 μM). One compound binds HIV-1 TAR with exceptional selectivity and inhibits Tat-mediated activation of the HIV-1 long terminal repeat by 81% in T cell lines and HIV replication in an HIV-1 indicator cell line (IC$_{50}$ ~23.1 μM).

INTRODUCTION

RNA is growing in its importance as a drug target[1] but current approaches used to identify protein-targeting small molecules are ill-suited for RNA. Most RNA targets lack the enzymatic activity required for conventional high throughput screening and their conformation switching activity proves difficult to assay experimentally[2–4]. Computational docking[5] can in principle provide the structural information needed to generally infer activity and can be used to screen uncharted regions of chemical space for novel RNA

binders[6]. However, current protocols fail to take into account the large conformational changes flexible RNA receptors undergo on binding small molecules, limiting discovery to compounds that target a narrow region of the structure landscape[7–9].

There is growing evidence that small molecules trigger RNA conformational changes by binding to conformers from pre-existing dynamic ensembles[10–17]. We recently introduced a general approach[11] for visualizing RNA dynamic ensembles, with the atomic resolution required for computational screening, and extended timescales (<milliseconds) needed to broadly sample the entire structure landscape. Here, multiple sets of NMR residual dipolar coupling (RDC) data that report on the dynamics of bond vectors relative to elongated RNA helices[15] are used to guide selection[18] of conformers from a large pool generated using MD[11]. By finding the minimum number of conformers that satisfy all time-averaged RDC data[19], a compact ensemble is constructed that samples unique and dominant positions across the entire RNA structure landscape. This unique combination of experiment and theory is critical for defining RNA ensembles given that the conformational space that has to be sampled is vast and difficult to reduce, and that current force fields remain underdeveloped for RNA compared to proteins. Using this approach, we constructed a dynamic ensemble for the transactivation response element (TAR)[12] from the human immunodeficiency type 1 (HIV-1) and showed that the twenty conformers in the TAR$^{\text{NMR-MD}}$ ensemble include structures that are very similar to structures of TAR observed when bound to seven distinct small molecules[11] (Fig. 1a and Supplementary Results Fig. 1).

Here, we show that docking small molecules onto RNA dynamic ensembles generated by NMR and MD provides a solution to the problem of taking into account large degrees of RNA conformational adaptation during virtual screening. Using this approach, we successfully identify six small molecules containing novel RNA binding moieties that bind TAR with near record affinity and inhibit its interaction with a Tat peptide *in vitro* (K_is ranging between 710 nM–169 μM). One of the compounds binds HIV-1 TAR with exceptional selectivity through unique interactions involving both the bulge and apical loop, and specifically inhibits Tat-mediated activation of the HIV-1 long terminal repeat by 81% in T cell lines and HIV replication in an HIV-1 indicator cell line (IC$_{50}$ of ~23.1 μM). From these studies, a new strategy emerges for selectively targeting highly flexible RNAs.

RESULTS

Accurate docking against known bound RNA structures

Docking small molecules onto individual conformers within NMR-MD ensembles, rather than a single static conformation, provides a natural but as of yet unrealized approach for taking into account RNA conformational adaptation during virtual screening. Such an approach inherently assumes that computational docking can be used to predict RNA-small molecule interactions, with sufficient accuracy, when the small molecule bound RNA structure is known. We therefore benchmarked binding predictions using the Internal Coordinate Mechanics (ICM, Molsoft LLC)[20] docking program for the ideal case in which the small molecule bound RNA structure is known. We used a diverse set of 96 small molecule-bound RNA structures, 48 of which had corresponding experimental K_ds. Structures with highly flexible small molecules (N$_{\text{flex}}$ >20; N$_{\text{flex}}$ is defined as the number of flexible torsions in the small molecule) that pose conformational sampling problems were excluded from analysis[21]. For each complex, the small molecule was removed, energy minimized in the absence of RNA, and re-docked onto the target RNA structure using fully flexible small molecule simulations. The RNA binding pocket, defined as all heavy atoms within 5 Å of the small molecule, was held rigid. In each case, the lowest docking score obtained from a specified number of iterations sampling different small molecule

conformations and poses was recorded (see Supplementary Methods). The binding energies were predicted with very good accuracy ($R = 0.71$) (Fig. 1b) and at a level comparable to state-of-the-art protein docking predictions[22]. More than half (53%) of the predicted binding poses match the X-ray/NMR structure to within a heavy atom root-mean-square-derivation (RMSD) cut-off of 2.5 Å (Supplementary Results Fig. 2a). This success rate compares well with the variability in the NMR bundle of structures, which typically results in an average RMSD of 1.8Å and in some cases >3Å[23]. Thus, the accuracy of docking predictions is not fundamentally limited by the scoring function or ability to sample different small molecule poses and conformations.

Overcoming 'adaptation' problem by docking RNA ensemble

A potentially more severe problem in RNA computational docking is that the small molecule bound RNA structure is generally not known, and can vary significantly from small molecule to small molecule. This uncertainty can relegate docking predictions into computational oblivion, particularly for highly flexible RNA receptors, which tend to undergo very large structural changes on binding small molecules. However, the impact of such uncertainty has never been quantified in RNA docking simulations. As an initial test, we examined how well computational docking could be used to predict the experimental binding energies for 38 TAR-binding compounds when docking against available X-ray[24] and NMR[25] structures of apo-TAR. Strikingly, the quality of the docking predictions deteriorates abruptly ($R = 0.13$) so as to become completely uninformative and ineffective in lead compound discovery (Supplementary Results Fig 2b). Docking against a computational (TAR^{MD}) ensemble consisting of 20 randomly chosen snap-shots from an 80 ns MD simulation of apo-TAR[11] resulted in some improvement ($R = 0.39$), but nowhere near the accuracy attainable when the bound RNA structure is known (Fig. 1b, $R = 0.71$). Here, each small molecule was independently docked onto each of the 20 conformers and the lowest overall score, corresponding to the dominant interaction energy, recorded. Thus, the accuracy of docking predictions is fundamentally limited by the uncertainty in the RNA bound structure.

We examined if one could recover the accuracy of docking predictions by docking small molecules against the 20 conformers in the TAR^{NMR-MD} ensemble. This NMR-informed ensemble was previously shown to sample many of the known ligand bound TAR conformations[11]. Remarkably, the binding energies are now predicted with an accuracy ($R = 0.66$) (Fig. 1c) that is comparable to that attained when the bound RNA structure is known ($R = 0.71$) (Fig. 1b). These results reinforce the view that small molecules do not 'induce' new TAR conformations, but rather, 'capture' conformers from a pre-existing dynamic ensemble and that TAR^{NMR-MD} provides a good approximation for this ensemble.

Virtual screening TAR dynamic ensemble

The interaction between TAR and the viral transactivator protein Tat has long been targeted for inhibiting HIV replication but has not yet resulted in clinically efficacious drugs[26]. We used our ensemble-targeted approach to identify TAR-targeting compounds. Each of the 20 conformers in the TAR^{NMR-MD} ensemble was subjected to virtual screening against ~51,000 small molecules (see **Methods**). The top 57 commercially available hits were experimentally tested using fluorescence-based assays (Supplementary Results Fig. 3 and 4) that probe (i) binding to a TAR construct containing a fluorescent probe at bulge residue U25[27] and (ii) inhibition of the interaction between TAR and an N-terminal-labeled-fluorescein peptide containing the arginine rich motif of TAR's cognate protein target Tat[28]. Six small molecules (Table 1) were experimentally validated in this manner that bind TAR ($K_d = 55$ nM–122μM) (Table 1) and that inhibit its interaction with Tat ($K_i = 710$ nM–169 μM) (Table 1). Together with spermine (Supplementary Results Fig. 5), the compounds are

identified with a hit rate of 12% and as high as 50% when only focusing on water-soluble compounds that did not require DMSO in experimental assays (DMSO was not included in docking simulations). This can be compared to a hit rate of 0% when screening 57 randomly selected small molecules from the same libraries (Supplementary Table 1). The virtual screen also identifies several small molecules, including two aminoglycosides, which despite their multiple positive charges, are correctly predicted to bind TAR with much weaker affinity as verified using fluorescence-based binding assays (Supplementary Results Fig. 6a).

The six small molecules include novel compounds that add to the chemical diversity of known TAR-binding small molecules (Table 1), including the recently developed cyclic peptides that bind TAR with high affinity and specificity[29,30]. For example, mitoxantrone (Table 1) a known RNA binder[31], binds TAR with an affinity ($K_d \sim 55$ nM), which among non-neamine derivatives is second only to one other small molecule, WM5 ($K_d \sim 19$ nM)[32] identified over two decades of research targeting TAR. The compound 5-(N,N)-Dimethylamiloride (DMA) ($K_d \sim 122$ μM) lacks cationic groups, contains a novel RNA binding scaffold consisting of a 5-chloropyrazin-2-amine core (Table 1), and targets a unique pocket within the TAR apical loop (see below). This is a rare example of a small molecule binding exclusively to an RNA apical loop. The molecules also include the four semi-synthetic aminoglycosides (Table 1), amikacin ($K_d \sim 1.5$ μM), butirosin A ($K_d \sim 4.8$ μM), netilimcin ($K_d \sim 1.4$ μM) and sisomicin ($K_d \sim 0.73$ μM) none of which have previously been shown to bind TAR.

Netilmicin binds TAR with high selectivity

The small molecules appear to have widely different specificities, as assayed using a competition experiment, in which the K_d is re-measured following the addition of 100-fold excess tRNA[33] (Supplementary Results Fig. 7). While we observe a significant deterioration in the binding affinities of mitoxatrone, amikacin, and sisomicin, (K_ds reduced by factors of 27, 7, and 3 respectively), consistent with non-specific binding to tRNA, we observed little to no changes in the K_ds for netilmicin and DMA, indicating that these compounds bind HIV-1 TAR with high specificity (Supplementary Results Fig. 7). Interestingly, a single ethyl group significantly reduces the binding specificity of sisomicin as compared to netilmicin in this assay.

As a more stringent test of specificity, we used fluorescence-based assays to measure the binding affinity of the small molecules (excluding butirosin A, which became commercially unavailable at the onset of these experiments) to a panel of three RNAs that more closely resemble the TAR hairpin. This includes an HIV-2 TAR variant (HIV-2) which differs from HIV-1 TAR by deletion of a single bulge residue, insertion of a G-U base-pair, and swapping of a G-C base-pair in the upper stem, the prokaryotic ribosomal A-site hairpin (A-site) and the HIV-1 rev response element hairpin (RRE), both of which are binding sites for a broad range of aminoglycosides[34,35] (Fig. 2a). The binding assays (Supplementary Results Fig. 8a) yielded specificity profiles for the various compounds that mirror those observed with tRNA (Supplementary Results Fig. 7a). Netilmicin showed the highest and exquisite levels of selectivity. It binds the closely related HIV-2 TAR with negligible affinity, and to A-site and RRE RNA with 35 and 86 fold reduced affinity, respectively (Fig. 2b, c and Supplementary Results Fig. 8a). As an even more stringent test, we measured the binding affinity of netilmicin to a TAR mutant that features a single cytosine bulge residue deletion and observed a 16-fold reduction in binding affinity (Supplementary Results Fig. 8a). Once again, sisomicin, show markedly reduced specificity compared to netilmicin and binds RRE with an affinity comparable to that of HIV-1 TAR (Fig. 2b, c). Interestingly, DMA which binds to the TAR apical loop (see below) showed as expected, strong selectivity against A-site and RRE but not HIV-2 TAR (Fig. 2b, c). All other small molecules bind at least one

other RNA with an affinity comparable to that of HIV-1 TAR (Supplementary Results Fig. 8a). We were able to confirm these specificity trends using competition assays analogous to those described for tRNA using these RNA constructs as competitors (Supplementary Results Fig. 7a).

Testing predicted binding modes using NMR

We tested the docking predicted TAR-small molecule binding modes with site-specific resolution using NMR chemical shift mapping experiments. Many of the small molecules are predicted to bind conformers within a contiguous region (conformers 12–15) of the TAR[NMR-MD] structure landscape (Fig. 3a) characterized by near coaxial alignment of the helices (inter-helical bend angle <12°), as observed for TAR when bound to Tat mimics. Correspondingly, all of the small molecules induced chemical shift perturbations characteristic of coaxial stacking of TAR helices as observed with Tat peptides and divalent ions[36] (e.g. U23 and C24, Fig. 3b and Supplementary Results Fig. 9). Interestingly, netilmicin, which shows the highest TAR binding specificity *in vitro*, is also predicted to bind conformers within the TAR ensemble with the highest specificity, with one conformer (18) accounting for 66% of the TAR population. The small molecules are predicted to bind TAR using distinct modes and to contact various combinations of residues in the bulge, upper stem, and apical loop (Fig. 3c) which form critical interactions with Tat, providing a structural basis for inhibiting the TAR-Tat interaction. Correspondingly, the six small molecules induce distinct chemical shift perturbations, particularly for residues predicted to be within the binding pocket (Fig. 3b, c and Supplementary Results Fig. 9). Significant perturbations (>0.1 ppm) are observed for 87% of TAR sites that are predicted to be within 5Å of the small molecule (Fig. 3c, red spheres). Perturbations outside this cut-off (Fig. 3c, green spheres) typically correspond to nearby flexible residues, which likely change conformation on binding the small molecule. Even detailed aspects of the predicted binding modes are supported in certain cases by the NMR data (Fig. 3b, c and Supplementary Methods) including unique stacking of mitoxantrone on G26, distinct binding modes for netilmicin and sisomicin that are mediated by contacts involving netilmicin's unique ethyl group, and binding of DMA to a unique pocket within the TAR apical loop. Unlike other aminoglycosides and their conjugated derivatives that have been shown to bind the TAR bulge and upper/lower stems[37–39], netilmicin, along with the other aminoglycosides uncovered here, interact with the apical loop in addition to the bulge and upper stem, while DMA provides a rare example of a small molecule that exclusively targets an RNA apical loop.

Netilmicin inhibits Tat activation of HIV-1 LTR

Of the five compounds tested (excluding butirosin A, which became commercially unavailable at the onset of these experiments), netilmicin, which binds TAR with the highest specificity *in vitro* (Fig. 2b), inhibited Tat-mediated activation of the HIV-1 promoter by ~81% when compared to the control (Fig. 4a) as assayed in live human T cells using a luciferase reporter construct transfected into Jurkat T cells. The other four compounds failed to show activity in this assay likely due to their much weaker binding specificity though we cannot rule out other effects, such as differences between full length Tat employed in transfection assays and Tat peptides used in the *in vitro* displacement assay. As a control, we repeated measurements of netilmicin activity upon addition of phorbol 12-myristate 13-acetate (PMA), which activates the HIV-1 LTR in a Tat independent manner. If netilmicin does indeed block Tat-mediated activation of the HIV-1 promoter through its interaction with TAR, then no inhibition should be observed upon PMA-mediated activation. Indeed, netilmicin did not inhibit PMA-mediated stimulation of the HIV-1 promoter, thus strongly ruling out off-target effects (Fig. 4a).

To further assess the specificity of netilmicin, we measured its inhibitory activity when using an HIV-2 TAR promoter containing the same HIV-2 TAR sequence used in the binding assays (Fig. 2b). According to *in vitro* binding data (Fig. 2b), netilmicin binds HIV-2 TAR with negligible affinity. If the mode of action of netilmicin involves binding to TAR, one would predict it would be far less effective at inhibiting the stimulation of the HIV-2 promoter. Indeed, netilmicin did not inhibit stimulation of the HIV-2 promoter, rather a small increase is observed (Fig. 4b). Moreover, no inhibition was observed when using HIV-1 Tat to stimulate the HIV-2 transcriptional promoter, indicating that netilmicin does not bind and inhibit HIV-1 Tat, but rather, affects Tat-mediated transactivation through its interaction with HIV-1 TAR (Supplementary Results Fig. 10). Thus, netilmicin specifically inhibits activation of the HIV-1 but not the closely related HIV-2 promoter in cellular assays.

Netilmicin inhibits HIV-1 replication

Remarkably, netilmicin not only inhibited Tat activation, but also HIV-1 replication, as assayed using an HIV-1 indicator cell line, TZM-bl and the HIV-1 NL4-3 isolate, which contains the same HIV-1 TAR sequence used in the *in vitro* studies. Addition of netilmicin to cells prior to and during infection resulted in a significant decrease in HIV replication, yielding an IC_{50} value (~23.1 μM) (Fig. 4c) that is strikingly similar to the value measured *in vitro* in the Tat displacement assay ($K_i = 14.1$ μM) (Table 1). The similarity of these inhibition constants further supports that netilmicin inhibits HIV replication by targeting TAR. This IC_{50} compares favorably with EC_{50s} (range between 0.7 μM and 30 μM) measured using the same HIV-1 NL4-3 strain for the most potent aminoglycoside derivatives that have been designed as Tat mimetics, including, for example, the aminoglycoside-arginine conjugate of neomycin (NeoR6 ~0.7 μM)[40]. We further corroborated the above results in an *in vivo* assay by infecting the HUT-78 T-cell line with HIV-1 NL4-3 in the presence of 100 μM netilmicin. Every three days, HIV-1 replication was assessed by measuring the amount of p24 antigen in the culture supernatants. Lower amounts of p24 were observed for samples treated with netilmicin compared to vehicle alone, with the largest difference observed on day 9 (Fig. 4d). Inhibition by netilmicin toxicity rather than TAR binding was ruled out since trypan blue staining for cellular viability showed no significant differences between the vehicle and netilmicin treated cells. Finally, our *in vitro* and cellular (gene reporter) assays show that despite its strong chemical similarity to netilmicin, sisomicin is far less effective at inhibiting Tat-mediated activation of the HIV-1 promoter most likely because it binds TAR with significantly reduced specificity. If netimilicin inhibits HIV-1 replication by inhibiting the TAR-Tat interaction, we would expect sisomicin to be a far less potent inhibitor. Indeed, the IC_{50} value (~157.1 μM) measured for sisomicin in the same HIV replication assay is ~ 7 fold higher than that measured for netilmicin (Fig 4c). These differences are significant considering that a two-fold increase in the IC_{50} is typically observed with resistant viruses. These data provide additional support that netilmicin inhibits HIV replication by selectively inhibiting the TAR-Tat interaction, unlike NeoR and other aminoglycoside-arginine conjugates, for which data suggest a different mode of inhibition involving the blocking of viral entry[40].

DISCUSSION

Netilmicin is the first experimentally validated RNA-targeting compound with *in vivo* activity to be identified using a virtual screen. It exhibits exquisite binding selectivity to HIV-1 TAR and this appears to be an important determinant of its activity. This high specificity is achieved in part by a single ethyl substituent on a key cationic amine group (Table 1). Comparison of the binding modes of netilmicin and sisomicin reveals that this modification alone significantly reduces netilmicin's binding affinity for tRNA

(Supplementary Results Fig. 7a) and RRE (Fig. 2b, c) without affecting its binding affinity for HIV-1 TAR. The alkyl group may stereo-chemically block access to the cationic group in the more rigid and less adaptable RNA binding pockets of tRNA and RRE but not in the more flexible and malleable HIV-1 TAR. Consistent with this notion, reducing the TAR flexibility by deleting a single cytosine bulge residue, which is not observed to make direct contacts with netilmicin, results in a 16-fold reduction in the netilmicin binding affinity (Supplementary Results Fig. 8a). Previous studies have shown that the flexible TAR is more capable of accommodating conformationally restrained small molecules as compared to the less malleable A-site[41]. Stereochemical crowding of key cationic groups on small molecules may well prove to be a general strategy for enhancing selectivity towards highly flexible RNAs.

The method developed here for targeting highly flexible RNA receptors can also be implemented to target other highly flexible targets, including intrinsically unfolded proteins implied in neurodegenerative diseases for which traditional structure-and assay-based approaches have thus far failed, and for which NMR informed computational dynamic ensembles are beginning to emerge[42]. Although five of the six newly identified TAR binders reported here have previously been shown to bind other RNAs, the compound library used in our virtual screen has been optimized for protein high throughput screening, and is not enriched with compounds that have favorable nucleic acid binding properties. Our virtual screen can be scaled up to include millions of compounds that have favorable nucleic acid binding and drug-like properties. This will provide a much needed route for efficiently screening new regions of chemical space in search for novel RNA-targeting lead compounds.

METHODS

Virtual Screening TAR dynamic ensemble

Virtual screening simulations were performed using ICM (molsoft LLC. La Jolla, CA)[20] employing 20 TAR^{NMR-MD} conformers and ~51,000 small molecules. TAR binding pockets were defined using the ICM PocketFinder module and the small molecule protonation states were computed over pH = 5.4–9.4 using ChemAxon© (www.chemaxon.com). The small molecule library (total 51,226 compounds) used in the virtual screening consisted of 49,166 compounds obtained from the Center of Chemical Genomics (CCG) at University of Michigan and 2060 compounds from the in-house library. Small molecules in the in-house library were extracted and drawn from published reports of verified RNA-binding small molecules by using ChemDraw (CambridgeSoft). Both libraries were saved in sdf file format. The virtual screening simulations started by docking small molecules with N_{flex} <20 using a thoroughness of 1, followed by a second round of screening top ~10% scoring small molecules using a thoroughness of 10. The top 57 small molecules (58 including spermine) were subject to experimental validation. The 57 small molecules were obtained from Maybridge, Chembridge, LKT Labs, and Sigma and with the exception of sisomicin (purity ≥ 80%) all were guaranteed to be ≥ 95% pure.

Fluorescence-based binding assay

The binding assay employed a TAR construct labeled with 2-aminopurine at bulge residue U25. An alternative TAR construct labeled with fluorescein at the same residue was used to measure binding of small molecules of which absorbance spectrum overlapped with fluorescence of 2-aminopurine. Both RNAs were purchased from Integrated DNA Technologies. RNA was annealed by heating at 95° C for 5 minutes followed by dilution (100 nM) into working buffer (10 mM phosphate, 20 mM NaCl, 0.1 mM EDTA, pH 6.8) and cooling on ice for 2 hours. Samples were pre-equilibrated for 5 minutes following

addition of small molecule. Fluorescence intensity measurements were collected using a Fluoromax-2 Fluorimeter at an excitation wavelength of 320 nm and emission wavelength of 390 nm for 2-aminopurine and an excitation wavelength of 485 nm and an emission wavelength of 520 nm for fluorescein. Fluorescence intensity measurements were recorded in triplicate and normalized to unbound-TAR.

Fluorescence-based TAR-Tat displacement assay

Fluorescence Polarization (FP)-based displacement assay employed an N-terminus fluorescein labeled Tat peptide (N-AAARKKRRQRRR-C, Genscript Corp.) and an *in vitro* synthesized elongated TAR. The elongated TAR was used to increase the dynamic range of FP measurements. The TAR (60nM) was incubated with varying concentrations of small molecules for 10 minutes followed by another 10-minute incubation with the addition of fluorescein-labeled Tat peptide (10nM). The FP buffer consisted of 50 mM Tris, 100 mM NaCl, and 0.01% nonidet-P40 at pH = 7.4. UV absorption spectrum was recorded for each small molecule tested to ensure no spectral overlap with fluorescein. Fluorescence Polarization (FP) measurements were collected in triplicate using 384 well plates read with a PHERAstar Plus plate reader (BMG LABTECH) and a 485 nm excitation wavelength and 520 nm detection wavelength optic module. IC_{50} and K_i values were calculated using the Prizm software (GraphPad Software Inc.). K_i values and corresponding errors reported in Table 1 are within the 95% confidence interval.

NMR experiments

All NMR experiments were performed at 25 °C using an Avance Bruker 600 MHz spectrometer equipped with a 5 mm triple-resonance cryogenic probe. NMR buffer consisted of 15 mM sodium phosphate, 0.1 mM EDTA, 25 mM NaCl, and 10% D_2O at pH ~6.4. NMR spectra were processed and analyzed using NMRPipe[43] and SPARKY 3[44].

Cellular assays

Transfection assays were performed by pretreating cells with small molecule, vehicle, or water as control 24 hours prior to transfection. Data shown in Fig. 4a, b and Supplementary Results Fig. 10 was normalized relative to Renilla luciferase activity, and represents the average of three independent transfections. Student's T-test, comparing Tat and netilmicin treatment (Tat + Net) to Tat and vehicle treatment (Tat) was used to obtain the p-values in Fig. 4a, b and Supplementary Results Fig. 10. The HIV-1 indicator cell line TZM-bl, which expresses luciferase upon HIV-1 infection, was plated in 96-well plates and treated with netilmicin, or water for a control, 24 hours prior to infection. The cells were then infected with HIV-1 NL4-3. 48 hours post-infection, luciferase activity was quantified as relative light units (RLU). Values were normalized and an IC_{50} value computed using non-linear regression. T-cell line Hut78 were infected with the HIV-1 isolate NL4-3 and HIV-1 replication assessed using p24 ELISA. Half of the media was harvested and replaced with uninfected Hut78 cells with or without 100 μM netilmicin at three day intervals.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Cooper TA, Wan L, Dreyfuss G. RNA and disease. Cell. 2009; 136:777–793. [PubMed: 19239895]

2. Parsons J, et al. Conformational inhibition of the hepatitis C virus internal ribosome entry site RNA. Nat Chem Biol. 2009; 5:823–825. [PubMed: 19767736]

3. Blount KF, Breaker RR. Riboswitches as antibacterial drug targets. Nat Biotechnol. 2006; 24:1558–1564. [PubMed: 17160062]

4. Thomas JR, Hergenrother PJ. Targeting RNA with small molecules. Chem Rev. 2008; 108:1171–1224. [PubMed: 18361529]

5. Kuntz ID. Structure-based strategies for drug design and discovery. Science. 1992; 257:1078–1082. [PubMed: 1509259]

6. Filikov AV, et al. Identification of ligands for RNA targets via structure-based virtual screening: HIV-1 TAR. J Comput-Aided Mol Des. 2000; 14:593–610. [PubMed: 10921774]

7. Hermann T. Rational ligand design for RNA: the role of static structure and conformational flexibility in target recognition. Biochimie. 2002; 84:869–875. [PubMed: 12458079]

8. Cruz JA, Westhof E. The dynamic landscapes of RNA architecture. Cell. 2009; 136:604–609. [PubMed: 19239882]

9. Fulle S, Gohlke H. Molecular recognition of RNA: challenges for modelling interactions and plasticity. J Mol Recognit. 2010; 23:220–231. [PubMed: 19941322]

10. Zhang Q, Sun X, Watt ED, Al-Hashimi HM. Resolving the motional modes that code for RNA adaptation. Science. 2006; 311:653–656. [PubMed: 16456078]

11. Frank AT, Stelzer AC, Al-Hashimi HM, Andricioaei I. Constructing RNA dynamical ensembles by combining MD and motionally decoupled NMR RDCs: new insights into RNA dynamics and adaptive ligand recognition. Nucleic Acids Res. 2009; 37:3670–3679. [PubMed: 19369218]

12. Puglisi JD, Tan R, Calnan BJ, Frankel AD, Williamson JR. Conformation of the TAR RNA-arginine complex by NMR spectroscopy. Science. 1992; 257:76–80. [PubMed: 1621097]

13. Williamson JR. Induced fit in RNA-protein recognition. Nature Struct Biol. 2000; 7:834–837. [PubMed: 11017187]

14. Leulliot N, Varani G. Current topics in RNA-protein recognition: Control of specificity and biological function through induced fit and conformational capture. Biochemistry. 2001; 40:7947–7956. [PubMed: 11434763]

15. Zhang Q, Stelzer AC, Fisher CK, Al-Hashimi HM. Visualizing spatially correlated dynamics that directs RNA conformational transitions. Nature. 2007; 450:1263–1267. [PubMed: 18097416]

16. Latham MP, Zimmermann GR, Pardi A. NMR chemical exchange as a probe for ligand-binding kinetics in a theophylline-binding RNA aptamer. J Am Chem Soc. 2009; 131:5052–5053. [PubMed: 19317486]

17. Vaiana AC, Sanbonmatsu KY. Stochastic gating and drug-ribosome interactions. J Mol Biol. 2009; 386:648–661. [PubMed: 19146858]

18. Chen Y, Campbell SL, Dokholyan NV. Deciphering protein dynamics from NMR data using explicit structure sampling and selection. Biophys J. 2007; 93:2300–2306. [PubMed: 17557784]

19. Clore GM, Schwieters CD. Amplitudes of protein backbone dynamics and correlated motions in a small alpha/beta protein: correspondence of dipolar coupling and heteronuclear relaxation measurements. Biochemistry. 2004; 43:10678–10691. [PubMed: 15311929]

20. Abagyan R, Totrov M, Kuznetsov D. ICM - A new method for protein modeling and design - applications to docking and structure prediction from the distorted native conformation. J of Comput Chem. 1994; 15:488–506.

21. Lang PT, et al. DOCK 6: Combining techniques to model RNA-small molecule complexes. RNA. 2009; 15:1219–1230. [PubMed: 19369428]

22. Cheng T, Li X, Li Y, Liu Z, Wang R. Comparative assessment of scoring functions on a diverse test set. J Chem Inf Model. 2009; 49:1079–1093. [PubMed: 19358517]

23. Guilbert C, James TL. Docking to RNA via root-mean-square-deviation-driven energy minimization with flexible ligands and flexible targets. J Chem Inf Model. 2008; 48:1257–1268. [PubMed: 18510306]

24. Ippolito JA, Steitz TA. A 1.3-angstrom resolution crystal structure of the HIV-1 trans- activation response region RNA stem reveals a metal ion- dependent bulge conformation. P Natl Acad Sci USA. 1998; 95:9819–9824.

25. Aboul-ela F, Karn J, Varani G. Structure of HIV-1 TAR RNA in the absence of ligands reveals a novel conformation of the trinucleotide bulge. Nucleic Acids Res. 1996; 24:3974–3981. [PubMed: 8918800]

26. Yang M. Discoveries of Tat-TAR interaction inhibitors for HIV-1. Curr Drug Targets Infect Disord. 2005; 5:433–444. [PubMed: 16535863]

27. Bradrick TD, Marino JP. Ligand-induced changes in 2-aminopurine fluorescence as a probe for small molecule binding to HIV-1 TAR RNA. RNA. 2004; 10:1459–1468. [PubMed: 15273324]

28. Matsumoto C, Hamasaki K, Mihara H, Ueno A. A high-throughput screening utilizing intramolecular fluorescence resonance energy transfer for the discovery of the molecules that bind HIV-1 TAR RNA specifically. Bioorg Med Chem Lett. 2000; 10:1857–1861. [PubMed: 10969985]

29. Davidson A, Patora-Komisarska K, Robinson JA, Varani G. Essential structural requirements for specific recognition of HIV TAR RNA by peptide mimetics of Tat protein. Nucleic Acids Res. 201010.1093/nar/gkq713

30. Davidson A, et al. Simultaneous recognition of HIV-1 TAR RNA bulge and loop sequences by cyclic peptide mimics of Tat protein. Proc Natl Acad Sci USA. 2009; 106:11931–11936. [PubMed: 19584251]

31. White RJ, Durr FE. Development of Mitoxantrone. Invest New Drugs. 1985; 3:85–93. [PubMed: 2410393]

32. Parolin C, et al. New anti-human immunodeficiency virus type 1 6-aminoquinolones: mechanism of action. Antimicrob Agents Chemother. 2003; 47:889–896. [PubMed: 12604517]

33. Blount KF, Tor Y, Hamasaki K, Ueno A. Using pyrene-labeled HIV-1 TAR to measure RNA- small molecule binding aminoglycoside antibiotics, neamine and its derivatives as potent inhibitors for the RNA-protein interactions derived from HIV-1 activators. Nucleic Acids Res. 2003; 31:5490–5500. [PubMed: 14500811]

34. DeJong ES, Chang CE, Gilson MK, Marino JP. Proflavine acts as a Rev inhibitor by targeting the high-affinity Rev binding site of the Rev responsive element of HIV-1. Biochemistry. 2003; 42:8035–8046. [PubMed: 12834355]

35. Kaul M, Barbieri CM, Pilch DS. Fluorescence-based approach for detecting and characterizing anti biotic-induced conformational changes in ribosomal RNA: Comparing aminoglycoside binding to prokaryotic and eukaryotic ribosomal RNA sequences. J Am Chem Soc. 2004; 126:3447–3453. [PubMed: 15025471]

36. Stelzer AC, Kratz JD, Zhang Q, Al-Hashimi HM. RNA dynamics by design: biasing ensembles towards the ligand-bound state. Angew Chem Int Ed. 2010; 49:5731–5733.

37. Lapidot A, Berchanski A, Borkow G. Insight into the mechanisms of aminoglycoside derivatives interaction with HIV-1 entry steps and viral gene transcription. FEBS J. 2008; 275:5236–5257. [PubMed: 18803669]

38. Lapidot A, Vijayabaskar V, Litovchick A, Yu JG, James TL. Structure-activity relationships of amino glyco side-arginine conjugates that bind HIV-1 RNAs as determined by fluorescence and NMR spectroscopy. FEBS Lett. 2004; 577:415–421. [PubMed: 15556620]

39. Faber C, Sticht H, Schweimer K, Rosch P. Structural rearrangements of HIV-1 Tat-responsive RNA upon binding of neomycin B. J Biol Chem. 2000; 275:20660–20666. [PubMed: 10747964]

40. Cabrera C, et al. Anti-HIV activity of a novel aminoglycoside-arginine conjugate. Antiviral Res. 2002; 53:1–8. [PubMed: 11684311]

41. Blount KF, Zhao F, Hermann T, Tor Y. Conformational constraint as a means for understanding RNA-aminoglycoside specificity. J Am Chem Soc. 2005; 127:9818–9829. [PubMed: 15998086]

42. Boehr DD, Nussinov R, Wright PE. The role of dynamic conformational ensembles in biomolecular recognition. Nat Chem Biol. 2009; 5:789–796. [PubMed: 19841628]

43. Delaglio F, et al. NMRPipe - a multidimensional spectral processing system based on unix pipes. J Biomol NMR. 1995; 6:277–293. [PubMed: 8520220]

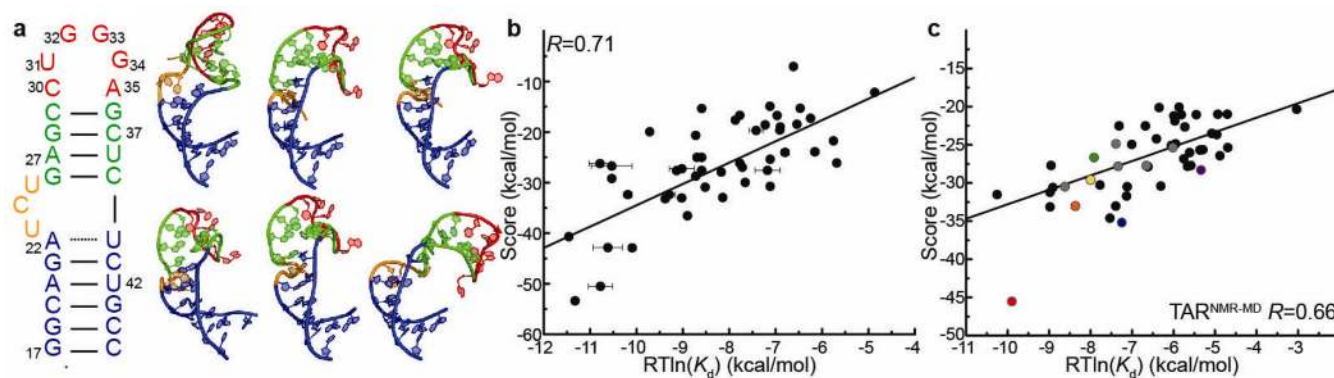44. Goddard, TD.; Kneller, DG. SPARKY. Vol. 3. University of California; San Franscisco:

Figure 1. Virtually screening RNA dynamic structure ensemble
a, Secondary structure of wild-type HIV-1 TAR RNA and representative conformers 3, 12–15, and 18 from the TAR[NMR-MD] dynamic ensemble[11]. **b–c** Correlation plots between experimental $\Delta G = RT\ln(K_d)$ with accompanying errors when reported, and ICM docking scores. Correlation coefficients (R) are shown in each case. Plots are shown for **b,** 48 RNA-small molecule complexes when docking small molecules onto the known small molecule bound RNA structure. **c,** 20 TAR conformers from the TAR[NMR-MD] ensemble.
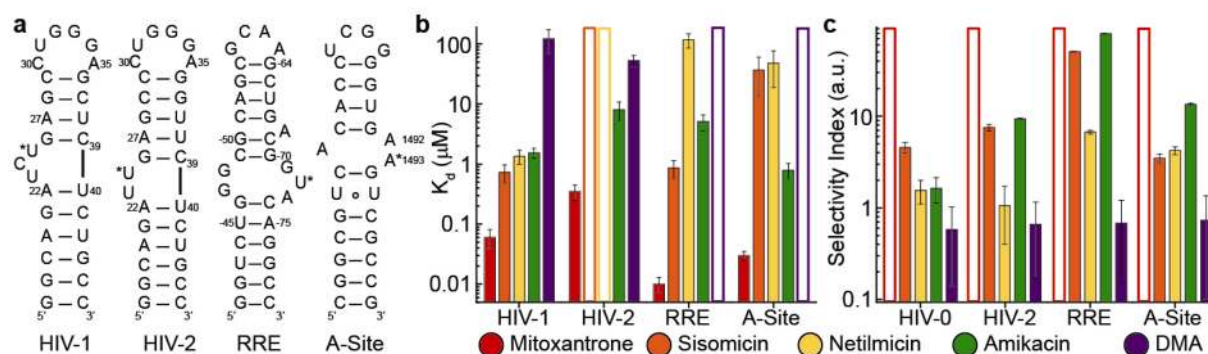
Figure 2. Analysis of small molecule binding specificity
a, Secondary structure of three related RNA hairpins (HIV-2 TAR, RRE, and the prokaryotic A-site) used in binding specificity analysis. Sites of 2′-fluorescein and 2-aminopurine base substitutions are indicated (*). **b,** Comparison of the dissociation constants (K_d) for each small molecule to each of four RNA constructs determined using fluorescence intensity measurements. Open symbols indicate a K_d could not be determined because the small molecule binds the RNA too weakly to saturate binding. **c,** Selectivity index (SI) defined as the ratio of the K_d measured for a small molecule to HIV-1 TAR in the presence and absence of competitor RNA indicated in the x-axis in parenthesis as the ratio between TAR and competitor. Open symbols indicate that a K_d could not be reliably determined due to strong competition from competitor RNA and incomplete HIV-1 TAR saturation. Note the apparent higher specificity of DMA can be attributed to binding to the apical loop and reduced susceptibility to competition by compounds that primarily target the bulge.
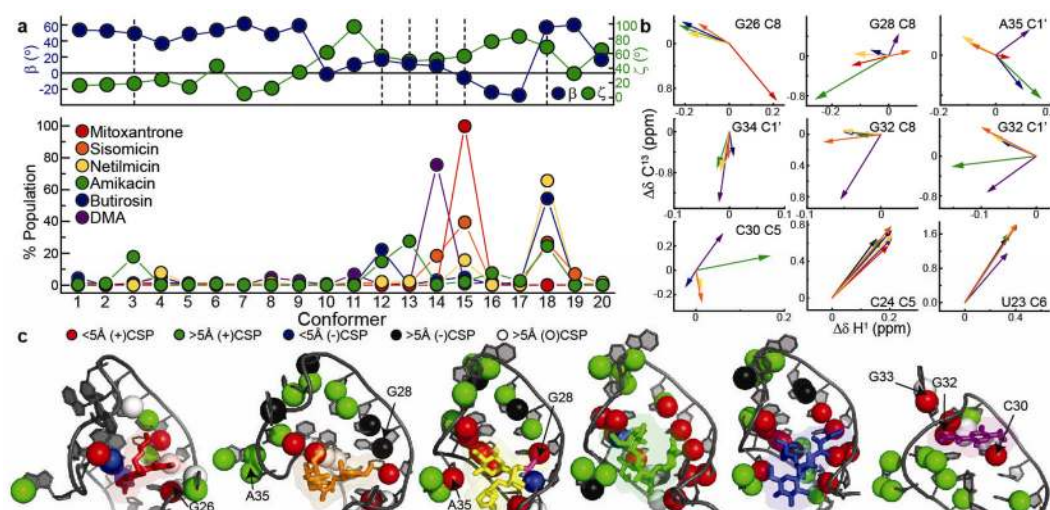
Figure 3. NMR site-specific characterization of TAR-small molecule binding modes
a, Docking predicted preferences for the binding of small molecules to distinct conformers in the TAR[NMR-MD] ensemble. Shown for each small molecule is the predicted fractional population bound to each of the 20 TAR conformers assuming the ICM computed energy. The inter-helical bend (β) and twist (ζ) angle is shown for each TAR conformer. **b,** Representative examples of chemical shift perturbation vectors colored according to the small molecule showing perturbations in resonance positions from free TAR to >97% small molecule bound TAR. Mg^{2+} perturbations are shown in black. NMR peaks are not shown for clarity and are provided in Supplementary Results Fig. 9. c, Mapping the NMR chemical shift perturbation data onto the docking predicted TAR-small molecule structures. TAR residues are colored according to whether they are predicted to be within or outside a 5Å distance cut-off from any atom in the small molecule and whether they exhibit significant (>0.1 ppm) chemical shift perturbations (+CSP) or not (−CSP), or if perturbations could not be assigned due to spectral overlap (○). The ethyl substituent of netilmicin is colored purple.
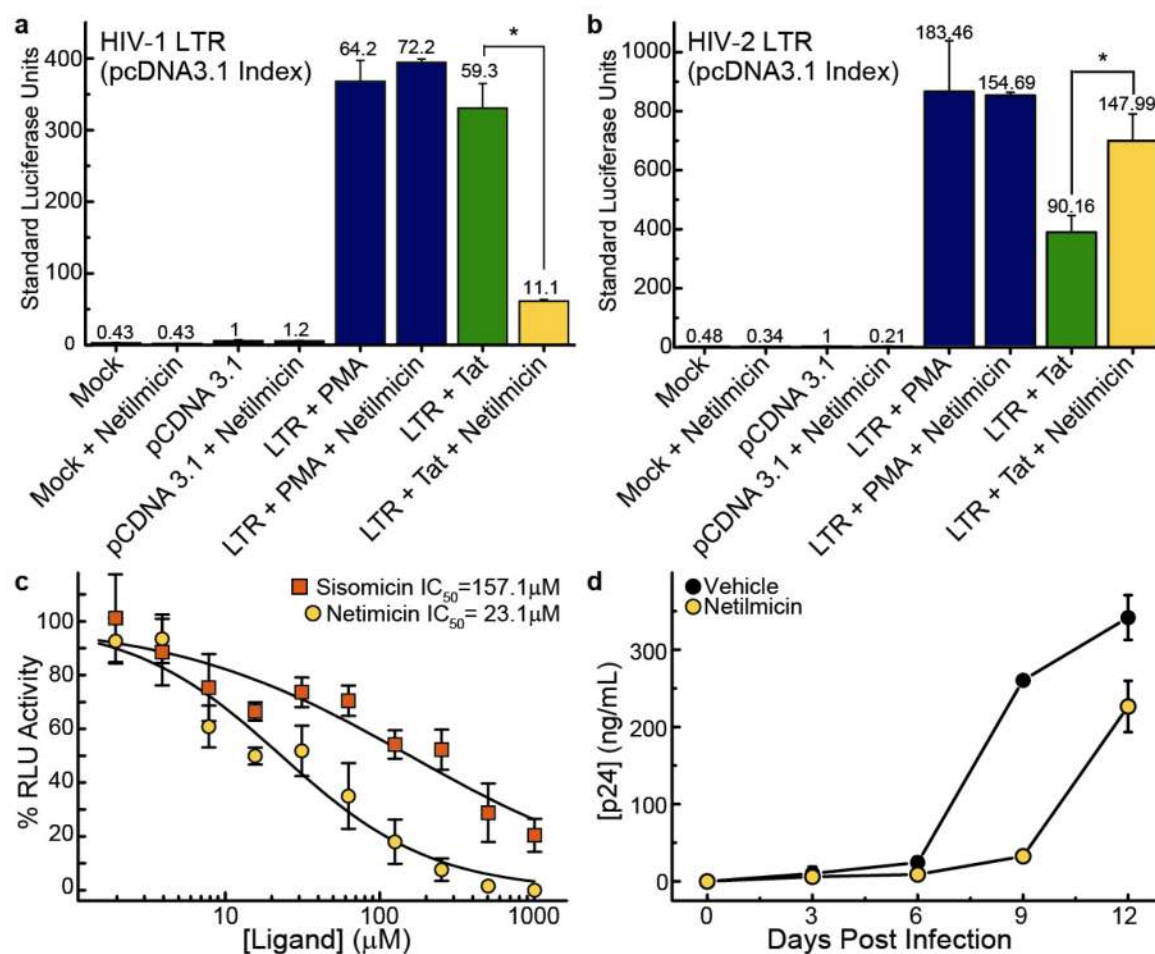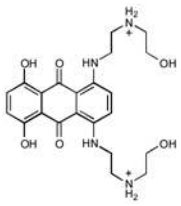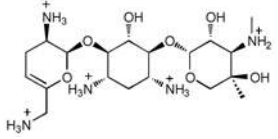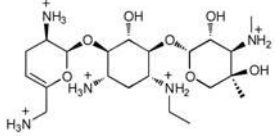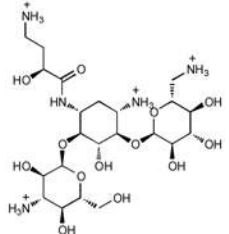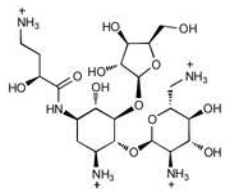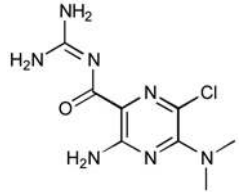
Figure 4. Netilmicin specifically inhibits Tat-mediated transactivation and HIV-1 replication
a, Jurkat T cells were treated with 100 μM netilmicin, or vehicle, for 24 hours and then transfected with the HIV-1 LTR-luc (LTR) and the HIV-1 Tat (Tat) constructs or empty vector (pcDNA3.1). "Mock" refers to the non-transfected control. Following transfection, netilmicin or vehicle was added to the media and luciferase activity measured 24 hours later. As a control, cells were transfected with only the HIV-1 LTR-luc and, 4 hours prior to the 24 hour completion point, were treated with 10 ng/mL PMA with or without netilmicin. Significance was calculated with *$p<0.02$. **b,** Experiments repeated using HIV-2 LTR with HIV-2 Tat. **c,** Netilmicin inhibits HIV-1 replication in the HIV-1 indicator cell line TZM-bl, which expresses luciferase upon HIV-1 infection. Error bars are obtained from duplicate measurements. For comparison, results when using sisomicin are also shown. **d,** Netilmicin inhibits HIV-1 replication in the T-cell line Hut78. Cells were infected with the HIV-1 isolate NL4-3. HIV-1 replication was assessed at three-day intervals by p24 ELISA. Error bars were calculated from duplicate measurements.

Table 1

Discovery of TAR binding small molecules by virtually screening its dynamic structure ensemble

Chemical structure of six small molecule hits along with dissociation constants (K_d) determined using fluorescence intensity measurements (mitoxantrone, sisomicin, netilmicin, amikacin, and 5-(N,N)-Dimethylamiloride) and NMR chemical shift perturbations (for butirosin A, due to commercial unavailability at the onset of these experiments) and inhibition constants (K_i) determined using fluorescence polarization measurements.

Structure	Compound	K_d (μM)	K_i (μM)
	Mitoxantrone (**1**)	0.055±0.021	0.71±0.32
	Sisomicin (**2**)	0.73±0.24	6.4±2.7
	Netilmicin (**3**)	1.35±0.35	14.1±8.3
	Amikacin (**4**)	1.54±0.29	16.2±9.6
	Butirosin A (**5**)	4.78±0.53	13.8±5.3
	5-(N,N)-Dimethyl amiloride (**6**)	121.85±50.65	169±98